Effect of Local Minima on Adiabatic Quantum Optimization

M. H. S. Amin
D-Wave Systems Inc., 100-4401 Still Creek Drive, Burnaby, B.C., V5C 6G9, Canada

We present a perturbative method to estimate the spectral gap for adiabatic quantum optimization, based on the structure of the energy levels in the problem Hamiltonian. We show that for problems that have exponentially large number of local minima close to the global minimum, the gap becomes exponentially small making the computation time exponentially long. The quantum advantage of adiabatic quantum computation may then be accessed only via the local adiabatic evolution, which requires phase coherence throughout the evolution and knowledge of the spectrum. Such problems, therefore, are not suitable for adiabatic quantum computation.

It is widely believed that quantum mechanics can provide speedup for certain computations. Different quantum algorithms have been proposed that potentially can solve problems such as factorization [1], unstructured search [2], or molecular simulations [3] on a quantum computer. One type of problems for which quantum mechanics may provide an advantage over classical computation is optimization. In optimization problems, one is interested in finding solutions that optimize some function subject to some constraints. Usually, not only the best solution, but also solutions close to it are of interest.

Physical systems at low temperatures naturally relax to their lowest energy states, effectively providing optimal solutions to their energy function. Such a relaxation process, however, may take a very long time. The time to reach the low energy states may be significantly reduced via an annealing process in which the temperature is reduced from a large value to a small value so slowly that the system stays effectively in equilibrium at all times. The slow evolution from a thermally disordered to ordered state with decreasing T will settle the system in one of its low lying energy states depending on the evolution time. Similar ideas have been employed in simulated annealing algorithms.

Quantum annealing (QA) [4, 5] is the quantum analog of the above classical annealing. In QA the disorder is introduced quantum mechanically via a Hamiltonian that does not commute with the optimization Hamiltonian. The added term generally has a ground state that is a superposition of all the eigenstates of the optimization Hamiltonian. Therefore, the disordered state is a superposition rather than a thermal mixture as it is in classical annealing. The disorder is removed by slowly removing the added term to the Hamiltonian. The system will then settle into one of its low lying energy states if the evolution time is long enough.

Closely related to QA is adiabatic quantum computation (AQC)[6]. In AQC an initial Hamiltonian $\mathcal{H}_B$ is slowly deformed into a final (problem) Hamiltonian $\mathcal{H}_P$:

$$\mathcal{H} = [1 - s(t)]\mathcal{H}_B + s(t)\mathcal{H}_P, \tag{1}$$

with $s(t)$ changing from 0 to 1 between the initial (t_i=0) and final (t_f) times. In this case, $\mathcal{H}_B$ plays the role of the disordering Hamiltonian. The main difference between QA and AQC is that in the latter, the system is constrained to its ground state at all times, starting from the ground state of $\mathcal{H}_B$, into which it is designed to be initialized, and ending in the ground state of $\mathcal{H}_P$, which encodes the solution to the problem of interest. In other words, AQC is an exact algorithm while QA is heuristic.

Unlike QA, AQC is not restricted to optimization problems, i.e., the problem Hamiltonian can be non-diagonal. For example, a universal AQC can run any quantum algorithm, and has been shown to be computationally equivalent to the gate model of quantum computation, as both can be efficiently mapped into each other [7, 8].

The performance of AQC is determined by the minimum gap g_m between the first two energy levels. In the *global* adiabatic evolution scheme, s is changed uniformly with time ($\dot{s}$ =const.) and the computation time depends on g_m as $\tau_{\rm global} \propto g_m^{-2}$. In the local adiabatic scheme [9], on the other hand, s is a nonlinear function of time designed in such a way to optimize the computation time by spending the majority of the evolution time in the vicinity of the anticrossing. As a result, the computation time of the local AQC is reduced to $\tau_{\rm local} \propto g_m^{-1}$, which scales better with g_m.

The global and local schemes of AQC are also different in terms of their response to decoherence. The global scheme is robust against environmental noise and decoherence [10, 11, 12, 13, 14]. The local adiabatic scheme, on the other hand, is very sensitive to decoherence. It was shown in Ref. [14] that in order for the local scheme to change the scaling of the computation time from $\propto g_m^{-2}$ to $\propto g_m^{-1}$, the computation time should be smaller than the global dephasing time. Moreover, local adiabatic evolution requires knowledge of the spectrum which is not feasible for general Hamiltonians.

An important question now is what kind of problems can benefit from the quantum advantage of AQC without requiring local adiabatic evolution and therefore phase coherence? It is known that for the unstructured search problem [9], $\tau_{\rm global} = O(N)$, which is the complexity of classical search, while $\tau_{\rm local} = O(\sqrt{N})$, which is the optimal performance of a quantum algorithm. (Here $N = 2^n$, where n is the number of qubits.) Thus, the advantage

over classical computation is only possible via the local adiabatic evolution. On the other hand, the universal AQC [7, 8, 15] can provide solution to a problem in polynomial time if the same problem can be solved in polynomial time using gate model quantum computation. Evidently, the polynomial advantage does not depend on local evolution, which can only provide quadratic enhancement. There have also been previous works to determine the complexity of AQC for some other special Hamiltonians [16, 17, 18]. In this letter, we study this problem for a rather general form of adiabatic quantum optimization.

We consider physically realizable initial and final Hamiltonians:

$$\mathcal{H}_B = -\frac{\Delta}{2}\sum_{i=1}^{n}\sigma_i^x, \tag{2}$$

$$\mathcal{H}_P = -\frac{\mathcal{E}}{2}\left[\sum_{i=1}^{n}h_i\sigma_i^z + \sum_{i,j=1}^{n}J_{ij}\sigma_i^z\sigma_j^z\right], \tag{3}$$

where $\sigma_i^{x,z}$ are the Pauli matrices for the i-th qubit, h_i and J_{ij} are dimensionless local fields and coupling coefficients respectively [typically $O(1)$], and $\mathcal{E}$ is some characteristic energy scale for $\mathcal{H}_P$. The initial Hamiltonian $\mathcal{H}_B$ has a nondegenerate ground state $|\psi_G(0)\rangle = |\overline{0^n}\rangle$. Here, we have adopted the notation $|\bar{z}\rangle = H^{\otimes n}|z\rangle$, $z \in \{0,1\}^n$, for states that are diagonal in the Hadamard basis, with H being the Hadamard transformation.

We denote the ground state of the total Hamiltonian by $|\psi_G\rangle = \sum_z a_z|z\rangle$, where a_z are complex probability amplitudes. At the beginning of the evolution, $a_z = 1/\sqrt{N}$, therefore $|\psi_G\rangle$ is a uniform superposition of all the states in the computation basis, but at the end of the evolution, it is only a superposition of the final solutions. The transition from large to small superpositions happens very suddenly at the minimum gap, which in the limit $g_m \to 0$ represents a quantum phase transition. Here, we only focus on first-order phase transition in which the gap is in the form of an avoided crossing [19]. If g_m is much smaller than the separation of the two crossing levels from other energy levels, then the slow evolution of the system close to the anticrossing will be restricted only to those levels. Using a new coordinate $\epsilon = 2E(s-s^*)$, where E is an energy scale characterizing the anticrossing and s^* is its position, one can write a two-state Hamiltonian:

$$\mathcal{H} = -(\epsilon\tau_z + g_m\tau_x)/2, \tag{4}$$

with $\tau_{x,z}$ being the Pauli matrices in the two-state subspace.

Immediately before and after the anticrossing, we write

$$|\psi_G^{\pm}\rangle = |\psi_G(\pm\epsilon_0)\rangle = \sum_{z\in\{0,1\}^n} a_z^{\pm}|z\rangle, \tag{5}$$

with $\epsilon_0 \ll E_{12}$, and E_{12} being the energy separation between the first two excited states. Using (4) it is easy to show that for $\epsilon_0 \gg g_m$

$$g_m \approx \epsilon_0|\langle\psi_G^+|\psi_G^-\rangle|, \tag{6}$$

i.e., g_m is proportional to the overlap of the wavefunctions before and after the anticrossing.

Let us introduce two sets

$$S^{\pm} = \{z:\ |a_z^{\pm}| > \delta\}, \tag{7}$$

where δ is a small number. Since all elements in $S^{\pm}$ contribute to the superposition, the normalization condition requires $|a_z^{\pm}| = O(1/\sqrt{|S^{\pm}|})$, yielding

$$|\psi_G^{\pm}\rangle \sim \sum_{z\in S^{\pm}}^{N}\frac{1}{\sqrt{|S^{\pm}|}}|z\rangle, \tag{8}$$

where $|S|$ denotes the cardinality of set S. The minimum gap will therefore be

$$g_m \propto \frac{|S^+\cap S^-|}{\sqrt{|S^+||S^-|}} \le \sqrt{\frac{|S^+|}{|S^-|}}. \tag{9}$$

The equality happens when all states in S^+ also belong to S^-. We shall only focus on this case as it provides an upper limit for g_m. In order to understand what can make the gap small, we need to understand how $S^{\pm}$ are constructed. To this end, we use perturbation expansion.

We introduce the dimensionless parameter

$$\zeta(t) \equiv \frac{s(t)\mathcal{E}}{[1-s(t)]\Delta}, \tag{10}$$

which varies from 0 to ∞ during the evolution. We shall drop the time dependence of ζ for simplicity. We start by calculating $|\psi_G\rangle$ near the end of the evolution, where ζ is large, by considering $\mathcal{H}_{1/\zeta} = \mathcal{H}_P + (1/\zeta)\mathcal{H}_B$. Let us for now assume that the problem has a unique solution, therefore the ground state of $\mathcal{H}_P$ is a nondegenerate state $|f\rangle$. To the 0-th order in $1/\zeta$, $|\psi_G^{(0)}\rangle = |f\rangle$. Since $\mathcal{H}_B$ is a linear function of σ_i^x, it can only generate single qubit flips. Thus, to include a state $|z\rangle$ with Hamming distance $m = ||z-f||$ from the solution $|f\rangle$, in $|\psi_G\rangle$, we need to apply $\mathcal{H}_B$ at least m times. Therefore, a_z^+ is nonzero only after the m-th order perturbation: $a_z^+ = O(1/\zeta^m)$. This restricts the states in S^+ to be close to f in Hamming distance. Requiring $1/\zeta^m \gtrsim \delta$, we find

$$S^+ \approx \{z:\ ||z-f|| < m_c\}, \tag{11}$$

$$m_c \propto \frac{\log(1/\delta)}{\log\zeta^+}, \qquad \zeta^{\pm} \equiv \zeta(\epsilon = \pm\epsilon_0). \tag{12}$$

The above argument can be easily generalized to multi-solution problems by writing the unperturbed ground state near the end of the evolution as $|\psi_G^{(0)}\rangle = N_s^{-1/2}\sum_{l=1}^{N_s}|f_l\rangle$, where f_l is the l-th solution

among the total N_s solutions of the problem. In this case,

$$S^+ \approx \{z : \min_l ||z - f_l|| < m_c\}. \quad (13)$$

Therefore, *the set S^+ is constructed from states that are close in Hamming distance to the final solutions.*

To find S^-, we perform perturbation expansion around $\mathcal{H}_B$, with ζ as the small parameter, using $\mathcal{H}_\zeta = \mathcal{H}_B + \zeta\mathcal{H}_P$. Before performing the perturbation expansion, let us use our intuition to understand how S^- can be formed. At $\zeta = 0$, the ground state of the Hamiltonian $\mathcal{H}_B$ is the uniform superposition of all the states in the computation basis. Adding a small perturbation $\zeta\mathcal{H}_P$ to the Hamiltonian will introduce a penalty to those eigenstates of $\mathcal{H}_P$ that have large eigenvalues. As a result, one expects that adding $\zeta\mathcal{H}_P$ will remove those high energy eigenstates from the superposition. The larger the ζ, the more high energy levels will be removed from the superposition and eventually only low lying states will survive.

Let us make this intuitive argument more quantitative. Since $\mathcal{H}_B$ is diagonal in the Hadamard basis, we need to do the perturbation expansion in that basis. Let us write

$$|\psi_G\rangle = \sum_{z\in\{0,1\}^n} b_{\bar{z}}|\bar{z}\rangle. \quad (14)$$

To the 0-th order, the wave function is $|\psi_G^{(0)}\rangle = |\overline{0^n}\rangle$. The Hamiltonian $\mathcal{H}_P$ is a bilinear function of σ_i^z, hence it can generate single and two qubit flips in the Hadamard basis. Again, in order to include a state $|\bar{z}\rangle$ into the superposition $|\psi_G\rangle$, where z has a Hamming weight w, we need to apply $\mathcal{H}_B$ at least $w/2$ times. This requires $b_{\bar{z}} = O(\zeta^{w/2})$ thereby making $b_{\bar{z}}$ non-negligible only if

$$w < w_c \propto \frac{\log(1/\delta)}{\log(1/\zeta^-)}. \quad (15)$$

In order to determine S^-, we need to know how $|\psi_G^-\rangle$ is formed in the computation basis, not in the Hadamard basis. Since $\mathcal{H}_B$ does not commute with $\mathcal{H}_P$, one can use the uncertainty principle to find the restriction imposed by the perturbation in the Hadamard basis, on the wave function in the computation basis. Let δE_P and δE_B represent uncertainties in $\mathcal{H}_P$ and $\mathcal{H}_B$, respectively:

$$\delta E_B \equiv (\langle\mathcal{H}_B^2\rangle - \langle\mathcal{H}_B\rangle^2)^{1/2},$$
$$\delta E_P \equiv (\langle\mathcal{H}_P^2\rangle - \langle\mathcal{H}_P\rangle^2)^{1/2}, \quad (16)$$

where $\langle ...\rangle$ represents expectation value. The uncertainty principle requires $\delta E_B \cdot \delta E_P \geq \frac{1}{2}\langle i[\mathcal{H}_B, \mathcal{H}_P]\rangle$. Every state $|\bar{z}\rangle$ is an eigenstate of $\mathcal{H}_B$ with eigenvalue $w\Delta$, where w is the Hamming weight of z. For the ground state $|\psi_G\rangle$, therefore, we have $\delta E_B \sim w_c\Delta$. The uncertainty principle requires $\delta E_P \propto 1/w_c$, leading to

$$S^- \approx \{z : E_z < E_c\}, \quad (17)$$

$$E_c \propto \frac{\log(1/\zeta^-)}{\log(1/\delta)}. \quad (18)$$

As expected, *the set S^- is made of low energy eigenstates of $\mathcal{H}_P$.*

Equations (12) and (18) suggest that as $\zeta^\pm \to 1$, $m_c \to \infty$ and $E_c \to 0$. Since the perturbation expansion breaks down at $\zeta \sim 1$, these equations cannot be extended all the way to $\zeta \sim 1$. In fact, $\zeta \sim 1$ is exactly where the quantum phase transition and therefore the anticrossing occurs. However, to calculate g_m using (6), we need $\epsilon_0 \gg g_m$, which ensures that $|\psi_G^\pm\rangle$ are indeed defined far away from the phase transition point, where the perturbation expansion and therefore (13) and (17) still hold. The important fact to notice is that the sets $S^\pm$ are formed in completely different ways: S^- is constructed by all the energy levels below some threshold, while S^+ is formed by all the states in Hamming proximity to the answers. The two sets could be very different leading to a very small energy gap.

For the upper limit in (9), i.e., $g_m \propto \sqrt{|S^+|/|S^-|}$, the computation time will be $\tau_{\rm local} \propto \sqrt{\tau_{\rm global}} \propto \sqrt{|S^-|/|S^+|}$. If the problem Hamiltonian happens to have an exponentially large number of low energy states that have large Hamming distances to the correct solutions (i.e., low energy *local minima*), those states will belong to S^- and not to S^+. The resulting gap will therefore be exponentially small, and the computation time will be extremely large. Especially, if $|S^-|$ becomes a fraction of N, then $\tau_{\rm global} = O(N)$, which is the complexity of the exhaustive search. The quantum advantage then will only be achievable via the local adiabatic scheme for which $\tau_{\rm local} = O(\sqrt{N})$.

An interesting example of such difficult problems is random 3-Satisfiability problem (3-SAT). These problems exhibit a phase transition when the ratio of the number of clauses m to the number of variables n reaches ≈ 4.2 [20]. Before the phase transition the number of solutions that satisfy the 3-SAT formula is extremely large, but suddenly after the phase transition point the number of satisfying solutions drops to zero. Therefore, near the phase transition point, by adding a few clauses to the 3-SAT formula, a large number of states that did satisfy it before will no longer do so. Those solutions, however, violate only those few clauses. In terms of energy, by adding a few terms in the Hamiltonian that provide penalties for those few clauses, an exponentially large number of states that used to be global minima suddenly become local minima but with energies very close to the ground state energy. This, therefore, would result in an exponentially small gap, as confirmed numerically [21] and analytically [22].

Another example is spin glasses [23], in which the local fields h_i are small or zero and coupling coefficients J_{ij} randomly couple (only) neighboring qubits. In such problems, domains can be formed if a large number of physically close qubits are strongly coupled to each other [24]. The qubits within a domain minimize the coupling terms

J_{ij} in $\mathcal{H}_P$. Those terms, however, remain unchanged if all of the qubits in the domain flip together. If the energy cost of flipping a domain, imposed by the field terms (h_i) and by the violation of the bounds (J_{ij}) at the domain boundary is not so large, then such a domain flipped state will form a low energy local minimum. If the size of the domain is large, then the Hamming distance between the local minimum and the global one will also be large. Thus, the local minimum and all the states close to it do not belong to S^+, while they do belong to S^-. A large number of such domains may make g_m exponentially small. It should be mentioned that if $h_i = 0$, then the final Hamiltonian will be symmetric under the total spin flip operation and the phase transition is likely to be second order, invalidating our assumption.

To conclude, we have used a perturbative approach to estimate the gap size for adiabatic quantum optimization problems. The gap is found to be inversely proportional to the square root of the number of states that have energies close to the global minimum. Therefore, problems that have a large number of low energy local minima tend to have a small gap. In general, problem instances in which the interaction terms in the Hamiltonian dominate the energy eigenvalues (i.e., typical values of J_{ij} are much larger than those for h_i) are likely to form low energy local minima and therefore make the gap small. If the number of low energy local minima becomes exponentially large, then the gap will be exponentially small. In such cases, only a local adiabatic evolution scheme can provide quantum advantage over classical computation. Local AQC, however, requires phase coherence during the evolution [14] and knowledge of the energy spectrum which limits its practicality. These problems, although unsuitable for AQC, could be suitable for heuristic algorithms (if approximate solutions are acceptable), because the chance of finding a solution within the acceptance threshold will be large. Quantum annealing therefore may provide good enough solutions in a short time, although finding the global minimum via AQC can take an extremely long time.

Finally, it should be mentioned that the energy gaps considered here are only the avoided crossing type, which correspond to first-order quantum phase transitions. If the final Hamiltonian possesses a symmetry that imposes a spontaneous symmetry breaking at the anticrossing, the resulting phase transition may become second order or higher orders. It has been stated that Hamiltonians with higher order phase transitions can provide better (possibly polynomial) scaling with the number of qubits [19, 25]. Study of those is beyond the scope of the present paper.

The author is grateful to A.J. Berkley, V. Choi, F. Chudak, R. Harris, W.M. Kaminsky, W.G. Macready, G. Rose, and C.J.S. Truncik for discussions.

[1] P. Shor, SIAM J. Comput. **26**, 1484 (1997).
[2] L. K. Grover, Phys. Rev. Lett. **79**, 325 (1997).
[3] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).
[4] J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli, Science **284**, 779 (1999).
[5] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, Science **295**, 2427 (2002).
[6] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).
[7] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, and S. Lloyd, SIAM J. Comput. **37**, 166 (2007).
[8] A. Mizel, D. A. Lidar, and M. Mitchell, Phys. Rev. Lett. **99**, 070502 (2007).
[9] J. Roland and N. J. Cerf, Phys. Rev. A **65**, 042308 (2002).
[10] A. M. Childs, E. Farhi, and J. Preskill, Phys. Rev. A **65**, 012322 (2001).
[11] J. Roland and N. J. Cerf, Phys. Rev. A **71**, 032330 (2005).
[12] M. H. S. Amin, P. J. Love, and C. J. S. Truncik, cond-mat/0609332.
[13] A. T. S. Wan, M. H. S. Amin, and S. X. Wang, cond-mat/0703085.
[14] M. H. S. Amin and D. V. Averin, arXiv:0708.0384.
[15] J. D. Biamonte and P. J. Love, arXiv:0704.1287.
[16] E. Farhi, J. Goldstone, S. Gutmann, and D. Nagaj, quant-ph/0512159.
[17] L. M. Ioannou and M. Mosca, quant-ph/0702241.
[18] W. van Dam, M. Mosca, and U. Vazirani, Proceedings of the 42nd Annual Symposium on Foundations of Computer Science, pp. 279-287 (2001).
[19] G. Schaller and R. Schützhold, arXiv:0708.1882.
[20] S. Kirkpatrick and B. Selman, Science **264**, 1297 (1994).
[21] M. Žnidarič, Phys. Rev. A **71**, 062305 (2005).
[22] M. Žnidarič and M. Horvat, Phys. Rev. A **73**, 022329 (2006).
[23] K. H. Fischer and J. A. Hertz, *Spin glasses* (Cambridge University Press, 1991).
[24] W. M. Kaminsky and S. Lloyd, in preparation.
[25] R. Schützhold and G. Schaller, Phys. Rev. A **74**, 060304 (2006).